Search Paper 12

# WEST Search History

DATE: Tuesday, July 08, 2003

| Set Name side by side | Query | Hit Count | Set Name result set |
|---|---|---|---|
| *DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=ADJ* | | | |
| L4 | holloway -james-l.in. | 0 | L4 |
| L3 | L2 not expressed | 38 | L3 |
| L2 | L1 same program | 190 | L2 |
| L1 | sequence same (match$ or align$) same tag | 1399 | L1 |

END OF SEARCH HISTORY

Search Page 12

Welcome to STN International! Enter x:x

LOGINID:SSSPTA1805JXB

PASSWORD:
TERMINAL (ENTER 1, 2, 3, OR ?):2

* * * * * * * * * * Welcome to STN International * * * * * * * * * *

| | | |
|---|---|---|
| NEWS 1 | | Web Page URLs for STN Seminar Schedule - N. America |
| NEWS 2 | | "Ask CAS" for self-help around the clock |
| NEWS 3 | Jun 03 | New e-mail delivery for search results now available |
| NEWS 4 | Aug 08 | PHARMAMarketLetter(PHARMAML) - new on STN |
| NEWS 5 | Aug 19 | Aquatic Toxicity Information Retrieval (AQUIRE) now available on STN |
| NEWS 6 | Aug 26 | Sequence searching in REGISTRY enhanced |
| NEWS 7 | Sep 03 | JAPIO has been reloaded and enhanced |
| NEWS 8 | Sep 16 | Experimental properties added to the REGISTRY file |
| NEWS 9 | Sep 16 | CA Section Thesaurus available in CAPLUS and CA |
| NEWS 10 | Oct 01 | CASREACT Enriched with Reactions from 1907 to 1985 |
| NEWS 11 | Oct 24 | BEILSTEIN adds new search fields |
| NEWS 12 | Oct 24 | Nutraceuticals International (NUTRACEUT) now available on STN |
| NEWS 13 | Nov 18 | DKILIT has been renamed APOLLIT |
| NEWS 14 | Nov 25 | More calculated properties added to REGISTRY |
| NEWS 15 | Dec 04 | CSA files on STN |
| NEWS 16 | Dec 17 | PCTFULL now covers WP/PCT Applications from 1978 to date |
| NEWS 17 | Dec 17 | TOXCENTER enhanced with additional content |
| NEWS 18 | Dec 17 | Adis Clinical Trials Insight now available on STN |
| NEWS 19 | Jan 29 | Simultaneous left and right truncation added to COMPENDEX, ENERGY, INSPEC |
| NEWS 20 | Feb 13 | CANCERLIT is no longer being updated |
| NEWS 21 | Feb 24 | METADEX enhancements |
| NEWS 22 | Feb 24 | PCTGEN now available on STN |
| NEWS 23 | Feb 24 | TEMA now available on STN |
| NEWS 24 | Feb 26 | NTIS now allows simultaneous left and right truncation |
| NEWS 25 | Feb 26 | PCTFULL now contains images |
| NEWS 26 | Mar 04 | SDI PACKAGE for monthly delivery of multifile SDI results |
| NEWS 27 | Mar 20 | EVENTLINE will be removed from STN |
| NEWS 28 | Mar 24 | PATDPAFULL now available on STN |
| NEWS 29 | Mar 24 | Additional information for trade-named substances without structures available in REGISTRY |
| NEWS 30 | Apr 11 | Display formats in DGENE enhanced |
| NEWS 31 | Apr 14 | MEDLINE Reload |
| NEWS 32 | Apr 17 | Polymer searching in REGISTRY enhanced |
| NEWS 33 | Jun 13 | Indexing from 1947 to 1956 added to records in CA/CAPLUS |
| NEWS 34 | Apr 21 | New current-awareness alert (SDI) frequency in WPIDS/WPINDEX/WPIX |
| NEWS 35 | Apr 28 | RDISCLOSURE now available on STN |
| NEWS 36 | May 05 | Pharmacokinetic information and systematic chemical names added to PHAR |
| NEWS 37 | May 15 | MEDLINE file segment of TOXCENTER reloaded |
| NEWS 38 | May 15 | Supporter information for ENCOMPPAT and ENCOMPLIT updated |
| NEWS 39 | May 16 | CHEMREACT will be removed from STN |
| NEWS 40 | May 19 | Simultaneous left and right truncation added to WSCA |
| NEWS 41 | May 19 | RAPRA enhanced with new search field, simultaneous left and right truncation |
| NEWS 42 | Jun 06 | Simultaneous left and right truncation added to CBNB |
| NEWS 43 | Jun 06 | PASCAL enhanced with additional data |
| NEWS 44 | Jun 20 | 2003 edition of the FSTA Thesaurus is now available |
| NEWS 45 | Jun 25 | HSDB has been reloaded |

```
 NEWS EXPRESS  April 4 CURRENT WINDOWS VERSION IS V6.01a, CURRENT
               MACINTOSH VERSION IS V6.0b(ENG) AND V6.0Jb(JP),
               AND CURRENT DISCOVER FILE IS DATED 01 APRIL 2003
 NEWS HOURS    STN Operating Hours Plus Help Desk Availability
 NEWS INTER    General Internet Information
 NEWS LOGIN    Welcome Banner and News Items
 NEWS PHONE    Direct Dial and Telecommunication Network Access to STN
 NEWS WWW      CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that
specific topic.

  All use of STN is subject to the provisions of the STN Customer
  agreement.  Please note that this agreement limits use to scientific
  research.  Use for software development or design or implementation
  of commercial gateways or other similar uses is prohibited and may
  result in loss of user privileges and other penalties.

  * * * * * * * * * * * * * * STN  Columbus  * * * * * * * * * * * * * * *

FILE 'HOME' ENTERED AT 15:24:57 ON 08 JUL 2003

=> file .pub
COST IN U.S. DOLLARS                                SINCE FILE      TOTAL
                                                         ENTRY    SESSION
FULL ESTIMATED COST                                       0.21       0.21

FILE 'MEDLINE' ENTERED AT 15:25:16 ON 08 JUL 2003

FILE 'BIOSIS' ENTERED AT 15:25:16 ON 08 JUL 2003
COPYRIGHT (C) 2003 BIOLOGICAL ABSTRACTS INC.(R)

=> s sequence and (match? or align?) and tag
L1         1648 SEQUENCE AND (MATCH? OR ALIGN?) AND TAG

=> s l1 and program
L2          115 L1 AND PROGRAM

=> duplicate remove l2
DUPLICATE PREFERENCE IS 'MEDLINE, BIOSIS'
KEEP DUPLICATES FROM MORE THAN ONE FILE?  Y/(N):n
PROCESSING COMPLETED FOR L2
L3            77 DUPLICATE REMOVE L2 (38 DUPLICATES REMOVED)

=> d 1-10 bib ab

L3   ANSWER 1 OF 77     MEDLINE
AN   2003248295     MEDLINE
DN   22656710   PubMed ID: 12771222
TI   Efficient clustering of large EST data sets on parallel computers.
AU   Kalyanaraman Anantharaman; Aluru Srinivas; Kothari Suresh; Brendel Volker
CS   Department of Computer Science, Iowa State University, Ames, IA 50011,
     USA.
SO   NUCLEIC ACIDS RESEARCH, (2003 Jun 1) 31 (11) 2963-74.
     Journal code: 0411011. ISSN: 1362-4962.
CY   England: United Kingdom
DT   (EVALUATION STUDIES)
     Journal; Article; (JOURNAL ARTICLE)
LA   English
FS   Priority Journals
EM   200306
ED   Entered STN: 20030529
     Last Updated on STN: 20030701
```

Entered Medline: 20030630

AB Clustering expressed **sequence tags** (ESTs) is a powerful strategy for gene identification, gene expression studies and identifying important genetic variations such as single nucleotide polymorphisms. To enable fast clustering of large-scale EST data, we developed PaCE (for Parallel Clustering of ESTs), a software **program** for EST clustering on parallel computers. In this paper, we report on the design and development of PaCE and its evaluation using Arabidopsis ESTs. The novel features of our approach include: (i) design of memory efficient algorithms to reduce the memory required to linear in the size of the input, (ii) a combination of algorithmic techniques to reduce the computational work without sacrificing the quality of clustering, and (iii) use of parallel processing to reduce run-time and facilitate clustering of larger data sets. Using a combination of these techniques, we report the clustering of 168 200 Arabidopsis ESTs in 15 min on an IBM xSeries cluster with 30 dual-processor nodes. We also clustered 327 632 rat ESTs in 47 min and 420 694 Triticum aestivum ESTs in 3 h and 15 min. We demonstrate the quality of our software using benchmark Arabidopsis EST data, and by comparing it with CAP3, a software widely used for EST assembly. Our software allows clustering of much larger EST data sets than is possible with current software. Because of its speed, it also facilitates multiple runs with different parameters, providing biologists a tool to better analyze EST **sequence** data. Using PaCE, we clustered EST data from 23 plant species and the results are available at the PlantGDB website.

L3 ANSWER 2 OF 77 MEDLINE

AN 2003311626 IN-PROCESS

DN 22723760 PubMed ID: 12840044

TI EST Mining and Functional Expression Assays Identify Extracellular Effector Proteins From the Plant Pathogen Phytophthora.

AU Torto Trudy A; Li Shuang; Styer Allison; Huitema Edgar; Testa Antonino; Gow Neil A R; Van West Pieter; Kamoun Sophien

CS Department of Plant Pathology, The Ohio State University, Ohio Agricultural Research and Development Center, Wooster, Ohio 44691, USA.

SO GENOME RESEARCH, (2003 Jul) 13 (7) 1675-85.
Journal code: 9518021. ISSN: 1088-9051.

CY United States

DT Journal; Article; (JOURNAL ARTICLE)

LA English

FS IN-PROCESS; NONINDEXED; Priority Journals

ED Entered STN: 20030704
Last Updated on STN: 20030704

AB Plant pathogenic microbes have the remarkable ability to manipulate biochemical, physiological, and morphological processes in their host plants.These manipulations are achieved through a diverse array of effector molecules that can either promote infection or trigger defense responses. We describe a general functional genomics approach aimed at identifying extracellular effector proteins from plant pathogenic microorganisms by combining data mining of expressed **sequence tags** (ESTs) with virus-based high-throughput functional expression assays in plants. PexFinder, an algorithm for automated identification of extracellular proteins from EST data sets, was developed and applied to 2147 ESTs from the oomycete plant pathogen Phytophthora infestans. The **program** identified 261 ESTs (12.2%) corresponding to a set of 142 nonredundant Pex (Phytophthora extracellular protein) cDNAs. Of these, 78 (55%) Pex cDNAs were novel with no significant **matches** in public databases. Validation of PexFinder was performed using proteomic analysis of secreted protein of P. infestans. To identify which of the Pex cDNAs encode effector proteins that manipulate plant processes, high-throughput functional expression assays in plants were performed on 63 of the identified cDNAs using an Agrobacterium tumefaciens binary vector carrying the potato virus X (PVX) genome. This led to the discovery of two novel necrosis-inducing cDNAs, crn1 and crn2, encoding extracellular proteins

that belong to a large and complex protein family in Phytophthora. Further characterization of the crn genes indicated that they are both expressed in P. infestans during colonization of the host plant tomato and that crn2 induced defense-response genes in tomato. Our results indicate that combining data mining using PexFinder with PVX-based functional assays can facilitate the discovery of novel pathogen effector proteins. In principle, this strategy can be applied to a variety of eukaryotic plant pathogens, including oomycetes, fungi, and nematodes.

L3 ANSWER 3 OF 77 MEDLINE
AN 2003278569 IN-PROCESS
DN 22690159 PubMed ID: 12805580
TI Refined annotation of the Arabidopsis genome by complete expressed **sequence tag** mapping.
AU Zhu Wei; Schlueter Shannon D; Brendel Volker
CS Department of Zoology and Genetics, Iowa State University, Ames, Iowa 50011-3260.
SO PLANT PHYSIOLOGY, (2003 Jun) 132 (2) 469-84.
Journal code: 0401224. ISSN: 0032-0889.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS IN-PROCESS; NONINDEXED; Priority Journals
ED Entered STN: 20030614
Last Updated on STN: 20030614
AB Expressed **sequence tags** (ESTs) currently encompass more entries in the public databases than any other form of **sequence** data. Thus, EST data sets provide a vast resource for gene identification and expression profiling. We have mapped the complete set of 176,915 publicly available Arabidopsis EST **sequences** onto the Arabidopsis genome using GeneSeqer, a spliced **alignment program** incorporating **sequence** similarity and splice site scoring. About 96% of the available ESTs could be properly **aligned** with a genomic locus, with the remaining ESTs deriving from organelle genomes and non-Arabidopsis sources or displaying insufficient **sequence** quality for **alignment**. The mapping provides verified sets of EST clusters for evaluation of EST clustering **programs**. Analysis of the spliced **alignments** suggests corrections to current gene structure annotation and provides examples of alternative and non-canonical pre-mRNA splicing. All results of this study were parsed into a database and are accessible via a flexible Web interface at http://www.plantgdb.org/AtGDB/.

L3 ANSWER 4 OF 77 MEDLINE DUPLICATE 1
AN 2003073130 MEDLINE
DN 22471709 PubMed ID: 12584131
TI Redundancy based detection of **sequence** polymorphisms in expressed **sequence tag** data using autoSNP.
AU Barker Gary; Batley Jacqueline; O' Sullivan Helen; Edwards Keith J; Edwards David
CS Institute of Arable Crop Research, Long Ashton, Bristol, BS41 9AF, UK.
SO BIOINFORMATICS, (2003 Feb 12) 19 (3) 421-2.
Journal code: 9808944. ISSN: 1367-4803.
CY England: United Kingdom
DT (EVALUATION STUDIES)
Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200306
ED Entered STN: 20030214
Last Updated on STN: 20030608
Entered Medline: 20030606
AB AutoSNP is a **program** to detect single nucleotide polymorphisms (SNPs) and insertion/deletion polymorphisms (indels) in expressed

**sequence tag** (EST) data. The **program** uses d2cluster and cap3 to cluster and **align** EST **sequences**, and uses redundancy to differentiate between candidate SNPs and **sequence** errors. Candidate polymorphisms are identified as occurring in multiple reads within an **alignment**. For each candidate SNP, two measures of confidence are calculated, the redundancy of the polymorphism at a SNP locus and the co segregation of the candidate SNP with other SNPs in the **alignment**. AVAILABILITY: The **program** was written in PERL and is freely available to non-commercial users by request from the authors.

L3 ANSWER 5 OF 77 MEDLINE
AN 2003056173 MEDLINE
DN 22453528 PubMed ID: 12566410
TI A complexity reduction algorithm for analysis and annotation of large genomic **sequences**.
AU Chuang Trees-Juen; Lin Wen-Chang; Lee Hurng-Chun; Wang Chi-Wei; Hsiao Keh-Lin; Wang Zi-Hao; Shieh Danny; Lin Simon C; Ch'ang Lan-Yang
CS Bioinformatics Research Center, Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan.
SO GENOME RESEARCH, (2003 Feb) 13 (2) 313-22.
Journal code: 9518021. ISSN: 1088-9051.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200303
ED Entered STN: 20030205
Last Updated on STN: 20030322
Entered Medline: 20030321
AB DNA is a universal language encrypted with biological instruction for life. In higher organisms, the genetic information is preserved predominantly in an organized exon/intron structure. When a gene is expressed, the exons are spliced together to form the transcript for protein synthesis. We have developed a complexity reduction algorithm for **sequence** analysis (CRASA) that enables direct **alignment** of cDNA **sequences** to the genome. This method features a progressive data structure in hierarchical orders to facilitate a fast and efficient search mechanism. CRASA implementation was tested with already annotated genomic **sequences** in two benchmark data sets and compared with 15 annotation **programs** (10 ab initio and 5 homology-based approaches) against the EST database. By the use of layered noise filters, the complexity of CRASA-**matched** data was reduced exponentially. The results from the benchmark tests showed that CRASA annotation excelled in both the sensitivity and specificity categories. When CRASA was applied to the analysis of human Chromosomes 21 and 22, an additional 83 potential genes were identified. With its large-scale processing capability, CRASA can be used as a robust tool for genome annotation with high accuracy by **matching** the EST **sequences** precisely to the genomic **sequences**.

L3 ANSWER 6 OF 77 MEDLINE
AN 2003178720 IN-PROCESS
DN 22583499 PubMed ID: 12697457
TI Cloning and function analysis of full-length cDNA **sequence** of Schistosoma japonicum eukaryotic translation initiation factor 2 alpha subunit.
AU Lu Xiao-Zhao; Peng Hong-Juan; Chen Xiao-Guang
CS Department of Parasitology, First Military Medical University, Guangzhou 510515, China.
SO Di Yi Jun Yi Da Xue Xue Bao, (2003 Apr) 23 (4) 296-9.
Journal code: 9426110. ISSN: 1000-2588.
CY China
DT Journal; Article; (JOURNAL ARTICLE)

LA English
FS IN-PROCESS; NONINDEXED; Priority Journals
ED Entered STN: 20030417
Last Updated on STN: 20030417
AB OBJECTIVE: To subclone the novel gene screened from the Schistosoma japonicum cercariae cDNA library through expressed **sequence tag** (EST) strategy and analyze its functions. METHOD: The cDNA fragment inserted in pTriplEx2 vector was sequenced and the result retrieved with BLASTn **program**. It was found that this cDNA was highly homologous to Schistosoma mansoni eukaryotic translation initiation factor 2 alpha subunit (eIF2alpha) mRNA. According to the known EST **sequence**, the 3'-terminal primer that **matched** the sequencing primer for the 5'-terminal in pTriplEx2 plasmid was designed and used to amplify the full-length open reading frame (ORF) **sequence** of the eIF2alpha from the cDNA Library. After proper purification, the PCR product was linked to pGEM-T vector and the recombinant T-vector was sequenced to obtain the full length ORF, which was retrieved for homologue identification using NCBI blast **program**. The **sequences** that were highly homologous underwent comparison at the levels of amino acids and nucleotides using BLAST 2 **Sequence program** on NCBI BLAST site. The motif and conserved domain were also retrieved with the software available online. RESULT: A novel cDNA **sequence** coding for a eIF2alpha was found from the cDNA library of Schistosoma japonicum cercariae, which was highly homologous to the known Schistosoma mansoni eIF2alpha mRNA, with the homology of 87% at the nucleotide level and 79% at the amino acid level. CONCLUSION: The novel gene found by EST strategy may encode a eIF2alpha which is highly homologous to Schistosoma mansoni eukaryotic eIF2alpha mRNA.

L3 ANSWER 7 OF 77 MEDLINE
AN 2003046423 MEDLINE
DN 22443406 PubMed ID: 12556150
TI Generation and characterization of cDNA clones from Sarcoptes scabiei var. hominis for an expressed **sequence tag** library: identification of homologues of house dust mite allergens.
AU Fischer Katja; Holt Deborah C; Harumal Pearly; Currie Bart J; Walton Shelley F; Kemp David J
CS The Queensland Institute of Medical Research, The Australian Centre for International and Tropical Health and Nutrition, and The University of Queensland, Brisbane, Australia.
SO AMERICAN JOURNAL OF TROPICAL MEDICINE AND HYGIENE, (2003 Jan) 68 (1) 61-4. Journal code: 0370507. ISSN: 0002-9637.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Abridged Index Medicus Journals; Priority Journals
OS GENBANK-AF317670; GENBANK-AF462195
EM 200302
ED Entered STN: 20030131
Last Updated on STN: 20030204
Entered Medline: 20030203
AB Molecular studies on scabies, a disease of considerable human and veterinary significance, have been limited because of the difficulty of obtaining the causative organism Sarcoptes scabiei, the "itch mite." We have used skin from the bedding of crusted scabies patients as a source of mites for the construction of libraries of cDNAs from S. scabiei var. hominis in the bacteriophage lambda vector lambdaZAP express. **Sequences** of 145 clones established that the libraries predominantly contain **sequences** from S. scabiei, enabling a major sequencing **program** to begin. Among those sequenced to date, cDNAs encoding S. scabiei homologues of 3 house dust mite allergens-the M-177 apolipoprotein, glutathione S-transferase, and paramyosin--were identified. The availability of cDNA libraries from S.

scabiei var. hominis and S. scabiei var. vulpes and the emerging public **sequence** databases from both opens up new possibilities in scabies research.

L3 ANSWER 8 OF 77 MEDLINE
AN 2002609959 MEDLINE
DN 22251143 PubMed ID: 12364589
TI Current methods of gene prediction, their strengths and weaknesses.
AU Mathe Catherine; Sagot Marie-France; Schiex Thomas; Rouze Pierre
CS Institut de Pharmacologie et Biologie Structurale, UMR 5089, 205 route de Narbonne, F-31077 Toulouse Cedex, France.. catherine.mathe@ipbs.fr
SO NUCLEIC ACIDS RESEARCH, (2002 Oct 1) 30 (19) 4103-17. Ref: 156
Journal code: 0411011. ISSN: 1362-4962.
CY England: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals
EM 200212
ED Entered STN: 20021008
Last Updated on STN: 20021218
Entered Medline: 20021216
AB While the genomes of many organisms have been sequenced over the last few years, transforming such raw **sequence** data into knowledge remains a hard task. A great number of prediction **programs** have been developed that try to address one part of this problem, which consists of locating the genes along a genome. This paper reviews the existing approaches to predicting genes in eukaryotic genomes and underlines their intrinsic advantages and limitations. The main mathematical models and computational algorithms adopted are also briefly described and the resulting software classified according to both the method and the type of evidence used. Finally, the several difficulties and pitfalls encountered by the **programs** are detailed, showing that improvements are needed and that new directions must be considered.

L3 ANSWER 9 OF 77 MEDLINE
AN 2002422531 MEDLINE
DN 22167277 PubMed ID: 12177459
TI Using genomic resources to guide research directions. The arabinogalactan protein gene family as a test case.
AU Schultz Carolyn J; Rumsewicz Michael P; Johnson Kim L; Jones Brian J; Gaspar Yolanda M; Bacic Antony
CS Department of Plant Science, Waite Agricultural Research Institute, The University of Adelaide, Glen Osmond, South Australia, Australia.. carolyn.schultz@adelaide.edu.au
SO PLANT PHYSIOLOGY, (2002 Aug) 129 (4) 1448-63.
Journal code: 0401224. ISSN: 0032-0889.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
OS PIR-AF497624
EM 200212
ED Entered STN: 20020815
Last Updated on STN: 20021217
Entered Medline: 20021203
AB Arabinogalactan proteins (AGPs) are extracellular hydroxyproline-rich proteoglycans implicated in plant growth and development. The protein backbones of AGPs are rich in proline/hydroxyproline, serine, alanine, and threonine. Most family members have less than 40% similarity; therefore, finding family members using Basic Local **Alignment** Search Tool searches is difficult. As part of our systematic analysis of AGP function in Arabidopsis, we wanted to make sure that we had identified most of the

members of the gene family. We used the biased amino acid composition of AGPs to identify AGPs and arabinogalactan (AG) peptides in the Arabidopsis genome. Different criteria were used to identify the fasciclin-like AGPs. In total, we have identified 13 classical AGPs, 10 AG-peptides, three basic AGPs that include a short lysine-rich region, and 21 fasciclin-like AGPs. To streamline the analysis of genomic resources to assist in the planning of targeted experimental approaches, we have adopted a flow chart to maximize the information that can be obtained about each gene. One of the key steps is the reformatting of the Arabidopsis Functional Genomics Consortium microarray data. This customized software **program** makes it possible to view the ratio data for all Arabidopsis Functional Genomics Consortium experiments and as many genes as desired in a single spreadsheet. The results for reciprocal experiments are grouped to simplify analysis and candidate AGPs involved in development or biotic and abiotic stress responses are readily identified. The microarray data support the suggestion that different AGPs have different functions.

L3 ANSWER 10 OF 77 MEDLINE DUPLICATE 2
AN 2002455196 MEDLINE
DN 22202135 PubMed ID: 12213779
TI GAZE: a generic framework for the integration of gene-prediction data by dynamic programming.
AU Howe Kevin L; Chothia Tom; Durbin Richard
CS The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.
SO GENOME RESEARCH, (2002 Sep) 12 (9) 1418-27.
Journal code: 9518021. ISSN: 1088-9051.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200210
ED Entered STN: 20020906
Last Updated on STN: 20021026
Entered Medline: 20021025
AB We describe a method (implemented in a **program**, GAZE) for assembling arbitrary evidence for individual gene components (features) into predictions of complete gene structures. Our system is generic in that both the features themselves, and the model of gene structure against which potential assemblies are validated and scored, are external to the system and supplied by the user. GAZE uses a dynamic programming algorithm to obtain the highest scoring gene structure according to the model and posterior probabilities that each input feature is part of a gene. A novel pruning strategy ensures that the algorithm has a run-time effectively linear in **sequence** length. To demonstrate the flexibility of our system in the incorporation of additional evidence into the gene prediction process, we show how it can be used to both represent nonstandard gene structures (in the form of trans-spliced genes in Caenorhabditis elegans), and make use of similarity information (in the form of Expressed **Sequence Tag alignments**), while requiring no change to the underlying software. GAZE is available at http://www.sanger.ac.uk/Software/analysis/GAZE.

=> d his

(FILE 'HOME' ENTERED AT 15:24:57 ON 08 JUL 2003)

FILE 'MEDLINE, BIOSIS' ENTERED AT 15:25:16 ON 08 JUL 2003
L1 1648 S SEQUENCE AND (MATCH? OR ALIGN?) AND TAG
L2 115 S L1 AND PROGRAM
L3 77 DUPLICATE REMOVE L2 (38 DUPLICATES REMOVED)

=> s l3 not expressed

L4 6 L3 NOT EXPRESSED

=> d 1-6 bib ab

L4 ANSWER 1 OF 6 MEDLINE
AN 2002028037 MEDLINE
DN 21378215 PubMed ID: 11485433
TI Comparison of energy-minimized structures of [Pd(II)(N-methyliminodiacetate)] complexes of X(1)-His-X(3)-His-His peptides as an analysis of steric and specific interactions with synthetic binding **tags** for IMAC separations.
AU Ward M S; Ataai M; Koepsel R R; Shepherd R E
CS Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA.
SO BIOTECHNOLOGY PROGRESS, (2001 Jul-Aug) 17 (4) 712-9.
Journal code: 8506292. ISSN: 8756-7938.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200112
ED Entered STN: 20020121
Last Updated on STN: 20020121
Entered Medline: 20011207
AB [Pd(II)(mida)(peptide)] complexes for the series of peptides of **sequence** X(1)-His-X(3)-His-His were studied by molecular mechanics methods using Spartan, MMFF94, and SYBYL **programs** with X(1) = X(3) = glycine (G), phenylalanine (F), tyrosine (Y), tryptophan (W), and with X(1) = glycine (G) and X(3) = proline (P). For comparison purposes, data were also obtained for the Ser-Pro-His-His-Gly (SPHHG) and the (His)(5) peptides.The latter two peptides and GHPHH are **tags** in current use for IMAC separations. These provide calibration points as to the binding affinities that have been determined for the entire series. The energies of the complexes, as an average trend found from the composite behavior of the three methods, were found to be SPHHG (205 kcal/mol) (most stable; are values obtained by MMFF94 methods) < HH(#)HH(#)H(#) (222; where # implies the site of attachment to **match** the other X(1)-His-X(3)-His-His peptides) < YHYHH (249) < GHGHH (265) < WHWHH (284) approximately GHPHH (286) < FHFHH (311) (least stable), implying that FHFHH might be a useful chromatographic **tag** for IMAC protein separations that would elute more readily than GHPHH from IMAC sites that are of square-planar structure, such as Cu(II)(ida-supported) IMAC columns. Specific H-bonded interactions are observed between the tyrosine X(1) and pendant carboxylates and between X(3) and the N-terminal amine of [Pd(mida)(YHYHH)]. Face-to-pi-face ring stacking occurs between phenylalanine X(1) and X(3) units in [Pd(mida)(FHFHH)], whereas edge C-H to pi H-bonding or pi stacking occurs between the X(1) and X(3) tryptophans of [Pd(mida)(WHWHH)]. Two energy minima were found with tryptophan. The more stable form has the aromatic rings more parallel, similar to the stacked form of phenylalanine, rather than the edge C-H to pi H-bonding, and virtually the same overall energy as for [Pd(mida)(GHPHH)]. The "perpendicular" structure was found as an initial local energy minimum, but additional MMFF94 calculations found the pi -stacked arrangement at energy ca. 39 kcal/mol lower than that of the nearly "perpendicular" arrangement of the tryptophan rings, a composite effect of relaxation of the peptide, together with differences in stabilities imparted by the differing geometries. The use of the terms "pi-stacked " and "perpendicular" forms represent the limiting cases available to the tryptophan side chain groups. A twist of about 15 degrees to 20 degrees in dihedral angle is all that is necessary to change between structures that are nearly described as one form or the other.

L4 ANSWER 2 OF 6 MEDLINE
AN 2000063248 MEDLINE

DN 20063248 PubMed ID: 10592214
TI tmRDB (tmRNA database).
AU Zwieb C; Wower J
CS Department of Molecular Biology, The University of Texas Health Science Center at Tyler, 11937 US Highway 271, Tyler, TX 75708-3154, USA.. zwieb@uthct.edu
NC GM-58267 (NIGMS)
SO NUCLEIC ACIDS RESEARCH, (2000 Jan 1) 28 (1) 169-70.
Journal code: 0411011. ISSN: 0305-1048.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200002
ED Entered STN: 20000314
Last Updated on STN: 20000314
Entered Medline: 20000225
AB The tmRNA database (tmRDB) is maintained at the University of Texas Health Science Center at Tyler, Texas, and is accessible on the WWW at URL http://psyche.uthct.edu/dbs/tmRDB/tmRDB.++ +html. A tmRDB mirror site is located on the campus of Auburn University, Auburn, Alabama, reachable at the URL http://www.ag.auburn.edu/mirror/tmRDB/. Since April 1997, the tmRDB has provided **sequences** of tmRNA (previously called 10Sa RNA), a molecule present in most bacteria and some organelles. This release adds 17 new **sequences** for a total of 60 tmRNAs. **Sequences** and corresponding tmRNA-encoded **tag** peptides are tabulated in alphabetical and phylo-genetic order. The updated tmRNA **alignment** improves the secondary structures of known tmRNAs on the level of individual basepairs. tmRDB also provides an introduction to tmRNA function in trans-translation (with links to relevant literature), a limited number of tmRNA secondary structure diagrams, and numerous three-dimensional models generated interactively with the **program** ERNA-3D.

L4 ANSWER 3 OF 6 MEDLINE
AN 1999352973 MEDLINE
DN 99352973 PubMed ID: 10424174
TI Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching.
AU Clauser K R; Baker P; Burlingame A L
CS Department of Pharmaceutical Chemistry, University of California, San Francisco 94143-0446, USA.
NC HD30367 (NICHD)
RR01614 (NCRR)
RR08282 (NCRR)
+
SO ANALYTICAL CHEMISTRY, (1999 Jul 15) 71 (14) 2871-82.
Journal code: 0370536. ISSN: 0003-2700.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199910
ED Entered STN: 19991026
Last Updated on STN: 19991026
Entered Medline: 19991008
AB We describe the impact of advances in mass measurement accuracy, +/- 10 ppm (internally calibrated), on protein identification experiments. This capability was brought about by delayed extraction techniques used in conjunction with matrix-assisted laser desorption ionization (MALDI) on a reflectron time-of-flight (TOF) mass spectrometer. This work explores the advantage of using accurate mass measurement (and thus constraint on the possible elemental composition of components in a protein digest) in strategies for searching protein, gene, and EST databases that employ (a)

mass values alone, (b) fragment-ion tagging derived from MS/MS spectra, and (c) de novo interpretation of MS/MS spectra. Significant improvement in the discriminating power of database searches has been found using only molecular weight values (i.e., measured mass) of > 10 peptide masses. When MALDI-TOF instruments are able to achieve the +/- 0.5-5 ppm mass accuracy necessary to distinguish peptide elemental compositions, it is possible to **match** homologous proteins having > 70% **sequence** identity to the protein being analyzed. The combination of a +/- 10 ppm measured parent mass of a single tryptic peptide and the near-complete amino acid (AA) composition information from immonium ions generated by MS/MS is capable of tagging a peptide in a database because only a few **sequence** permutations > 11 AA's in length for an AA composition can ever be found in a proteome. De novo interpretation of peptide MS/MS spectra may be accomplished by altering our MS-**Tag** **program** to replace an entire database with calculation of only the **sequence** permutations possible from the accurate parent mass and immonium ion limited AA compositions. A hybrid strategy is employed using de novo MS/MS interpretation followed by text-based **sequence** similarity searching of a database.

L4 ANSWER 4 OF 6 MEDLINE
AN 1999063679 MEDLINE
DN 99063679 PubMed ID: 9847168
TI The tmRNA database (tmRDB).
AU Wower J; Zwieb C
CS Department of Animal and Dairy Sciences, Auburn University, Auburn, AL 36849-5415, USA.
NC GM-58267 (NIGMS)
SO NUCLEIC ACIDS RESEARCH, (1999 Jan 1) 27 (1) 167.
Journal code: 0411011. ISSN: 0305-1048.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199903
ED Entered STN: 19990326
Last Updated on STN: 19990326
Entered Medline: 19990316
AB As of September, 1998, a total of 43 **sequences** are contained within the tmRNA database (tmRDB). The tmRNA **sequences** are arranged alphabetically and ordered phylogenetically. The **alignment** of the tmRNAs emphasizes the basepairs that are supported by comparative **sequence** analysis and establishes minimal secondary structures for the known tmRNAs. A corresponding **alignment** of the predicted tmRNA-encoded **tag** peptides is presented. The tmRDB also offers a small number of RNA secondary structure diagrams and PDB-formatted three-dimensional models generated with the **program** ERNA-3D. The data are available freely at the URL http://psyche.uthct.edu/dbs/tmRDB/tmRDB.++ +html

L4 ANSWER 5 OF 6 MEDLINE
AN 96205323 MEDLINE
DN 96205323 PubMed ID: 8630008
TI Rapid protein identification using N-terminal "**sequence** **tag**" and amino acid analysis.
AU Wilkins M R; Ou K; Appel R D; Sanchez J C; Yan J X; Golaz O; Farnsworth V; Cartier P; Hochstrasser D F; Williams K L; Gooley A A
CS Macquarie University Centre for Analytical Biotechnology, Sydney, New South Wales, Australia.
SO BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, (1996 Apr 25) 221 (3) 609-13.
Journal code: 0372516. ISSN: 0006-291X.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)

LA English
FS Priority Journals
EM 199606
ED Entered STN: 19960708
Last Updated on STN: 19960708
Entered Medline: 19960621
AB Proteins can be identified by amino acid analysis and database **matching**, but it is often desirable to increase the confidence in identity through the use of other techniques. Here we describe a rapid protein identification method that uses Edman degradation to create a 3 or 4 amino acid N-terminal "**sequence tag**," following which proteins are subjected to amino acid analysis protein identification procedures. Edman degradation methods have been modified to take only 23 min per cycle, and rapid amino acid analysis techniques are used. The Edman degradation and amino acid analysis is done on a single PVDF membrane-bound protein sample. A computer database **matching program** is also presented which uses both amino acid composition and "**sequence tag**" data for protein identification. This method represents the most inexpensive, accurate, and rapid means of protein identification, which is ideal for the screening of proteomes separated by 2-D gel electrophoresis. The creation of N-terminal Edman degradation "**sequence-tags**" prior to peptide mass fingerprinting of samples should also be useful.

L4 ANSWER 6 OF 6 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
AN 2001:443319 BIOSIS
DN PREV200100443319
TI BLAST OFF! Magnaporthe grisea genome project successfully launched.
AU Dean, R. A. (1)
CS (1) Fungal Genomics Laboratory, NC State University, Raleigh, NC, 27695 USA
SO Phytopathology, (June, 2001) Vol. 91, No. 6 Supplement, pp. S158. print.
Meeting Info.: Joint Meeting of the American Phytopathological Society, the Mycological Society of America, and the Society of Nematologists Salt Lake City, Utah, USA August 25-29, 2001
ISSN: 0031-949X.
DT Conference
LA English
SL English
AB Rice blast disease, caused by Magnaporthe grisea, is one of the most devastating threats to food security worldwide. The fungus is amenable to classical and molecular genetic manipulation and is a compelling experimental system for elucidating numerous aspects of pathogenesis, including infection-related morphogenesis, host species and cultivar specificity, and signaling pathways. In 1998, an international consortium (IRBGP) was established to **sequence** the rice blast genome. For this initiative, we used a 25X large insert (130 kb) HindIII BAC library to construct a physical map of the genome. BAC clones were fingerprinted and assembled into 188 contigs. These were **aligned** into a physical map by anchoring to mapped RFLP markers. Chromosome 7 (4.2 Mb) has been studied in the greatest detail and a set of 42 BAC clones representing a minimum tiling path covering >95% of the chromosome has been deduced. The entire BAC library was end sequenced providing **sequence tag** connectors (STC) every 3-4 kb across the genome. A federated database integrating physical, genetic and expression data from relational and object-oriented databases is being developed. We have initiated a draft **sequence** (apprx5X coverage) of chromosome 7 using the "BAC by BAC" approach coupled with information from our STC/fingerprint databases. A comprehensive EST **program** has been launched. 30,000 ESTs will be derived from 8 cDNA libraries prepared from different stages of growth and development as well as cells subjected to various stress conditions. A set of apprx5,000 ESTs representing unique genes will be further sequenced. The current status of the genome project and database development will be presented.

```
=> d his

     (FILE 'HOME' ENTERED AT 15:24:57 ON 08 JUL 2003)

     FILE 'MEDLINE, BIOSIS' ENTERED AT 15:25:16 ON 08 JUL 2003
L1          1648 S SEQUENCE AND (MATCH? OR ALIGN?) AND TAG
L2           115 S L1 AND PROGRAM
L3            77 DUPLICATE REMOVE L2 (38 DUPLICATES REMOVED)
L4             6 S L3 NOT EXPRESSED

=> s sequence and nieve
L5             0 SEQUENCE AND NIEVE

=> e holloway
E1             4     HOLLOTYPE/BI
E2         11848     HOLLOW/BI
E3           146 --> HOLLOWAY/BI
E4             1     HOLLOWAYAE/BI
E5             1     HOLLOWAYANUS/BI
E6             1     HOLLOWAYCRINUS/BI
E7             1     HOLLOWAYENSIS/BI
E8            28     HOLLOWAYI/BI
E9             1     HOLLOWAYORUM/BI
E10            1     HOLLOWAYS/BI
E11            2     HOLLOWCORE/BI
E12            1     HOLLOWDAY/BI

=> e holloway/au
E13            2     HOLLOWATY E/AU
E14            1     HOLLOWATY ERIC/AU
E15            0 --> HOLLOWAY/AU
E16           32     HOLLOWAY A/AU
E17           29     HOLLOWAY A C/AU
E18           71     HOLLOWAY A F/AU
E19           10     HOLLOWAY A J/AU
E20            1     HOLLOWAY A JR/AU
E21            6     HOLLOWAY A K/AU
E22           31     HOLLOWAY A M/AU
E23            4     HOLLOWAY A W/AU
E24            2     HOLLOWAY A Z/AU

=> e holloway J L/au
E25           26     HOLLOWAY J J/AU
E26            1     HOLLOWAY J J H FULLER/AU
E27           10 --> HOLLOWAY J L/AU
E28           22     HOLLOWAY J M/AU
E29            1     HOLLOWAY J N/AU
E30            2     HOLLOWAY J P/AU
E31            3     HOLLOWAY J R/AU
E32           12     HOLLOWAY J S/AU
E33           13     HOLLOWAY J T/AU
E34            1     HOLLOWAY J TED/AU
E35          133     HOLLOWAY J W/AU
E36            1     HOLLOWAY JACQUELINE A/AU

=> s e27
L6            10 "HOLLOWAY J L"/AU

=> duplicate remove l6
DUPLICATE PREFERENCE IS 'MEDLINE, BIOSIS'
KEEP DUPLICATES FROM MORE THAN ONE FILE?  Y/(N):n
PROCESSING COMPLETED FOR L6
```

L7 8 DUPLICATE REMOVE L6 (2 DUPLICATES REMOVED)

=> d 1-8 bib ab

L7 ANSWER 1 OF 8 MEDLINE
AN 2001393626 MEDLINE
DN 21124966 PubMed ID: 11223802
TI Social interaction with an intoxicated sibling can result in increased intake of ethanol by periadolescent rats.
AU Hunt P S; **Holloway J L**; Scordalakes E M
CS Department of Psychology, College of William & Mary, Williamsburg, VA 23187-8795, USA.. pshunt@wm.edu
NC AA12135 (NIAAA)
AA12466 (NIAAA)
SO DEVELOPMENTAL PSYCHOBIOLOGY, (2001 Mar) 38 (2) 101-9.
Journal code: 0164074. ISSN: 0012-1630.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200107
ED Entered STN: 20010716
Last Updated on STN: 20010716
Entered Medline: 20010712
AB A novel procedure for enhancing voluntary intake of ethanol in periadolescent rats is described. The procedure is a modification of Galef et al.'s (e.g., Galef, Kennett, & Stein, 1985; Anim Learn Behave 13:25-30) demonstrator-observer procedure. Subjects were Sprague-Dawley rats, 28-35 days of age. The experimental subject (observer) interacted with a same-sex conspecific (demonstrator) previously administered (a) 1.5 g/kg ethanol, (b) an equal volume of water, or (c) 2.1% Sanka coffee intragastrically. Observers were tested with 24-hour access to ethanol and coffee solutions. Observers that had interacted with demonstrators administered ethanol ingested significantly more ethanol during the test than observers in the other two groups. In Experiment 2 demonstrators were administered one of several doses of ethanol (0.0, 1.0, 1.5, or 3.0 g/kg) and observers' ethanol intakes were assessed. Only those observers that interacted with 1.5 g/kg demonstrators increased their ingestion of ethanol, relative to water controls. The lower (1.0 g/kg) and higher (3.0 g/kg) dose groups did not show altered ethanol ingestion. These results are discussed with respect to threshold levels of respired ethanol cues and the ability of observers to detect these cues from demonstrators. The demonstrator-observer procedure appears to be effective for the social transmission of preferences for ethanol in periadolescent rats.
Copyright 2001 John Wiley & Sons, Inc.

L7 ANSWER 2 OF 8 MEDLINE
AN 2000179689 MEDLINE
DN 20179689 PubMed ID: 10713450
TI Molecular cloning, chromosome mapping and characterization of a testis-specific cystatin-like cDNA, cystatin T.
AU Shoemaker K; **Holloway J L**; Whitmore T E; Maurer M; Feldhaus A L
CS Department of Genetics, ZymoGenetics Inc., Seattle, WA, USA.
SO GENE, (2000 Mar 7) 245 (1) 103-8.
Journal code: 7706761. ISSN: 0378-1119.
CY Netherlands
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200005
ED Entered STN: 20000512
Last Updated on STN: 20000512
Entered Medline: 20000504
AB The cystatin superfamily of cysteine proteinase inhibitors consists of

three major families. In the present study, we report the cloning of the cDNA for mouse cystatin T, which is related to family 2 cystatins. The deduced amino acid sequence of cystatin T contains regions of significant sequence homology including the four highly conserved cysteine residues in exact alignment with all cystatin family 2 members. However, cystatin T lacks some of the conserved motifs believed to be important for inhibition of cysteine proteinase activity. These characteristics are seen in two other recently cloned genes, CRES and Testatin. Thus, cystatin T appears to be the third member of the CRES/Testatin subgroup of family 2 cystatins. The mouse cystatin T gene was mapped on a region of chromosome 2 that contains a cluster of cystatin genes, including cystatin C and CRES. Northern blot analysis demonstrated that expression of mouse cystatin T is highly restricted to the mouse testis. Thus, a shared characteristic of the cystatin family 2 subgroup members is an expression pattern limited primarily to the male reproductive tract.

L7 ANSWER 3 OF 8 MEDLINE DUPLICATE 1
AN 2000511436 MEDLINE
DN 20515579 PubMed ID: 11060443
TI Human secretin (SCT): gene structure, chromosome location, and distribution of mRNA.
AU Whitmore T E; **Holloway J L**; Lofton-Day C E; Maurer M F; Chen L; Quinton T J; Vincent J B; Scherer S W; Lok S
CS ZymoGenetics Inc., Seattle, WA 98102, USA.
SO CYTOGENETICS AND CELL GENETICS, (2000) 90 (1-2) 47-52.
Journal code: 0367735. ISSN: 0301-0171.
CY Switzerland
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
OS GENBANK-AF244355
EM 200011
ED Entered STN: 20010322
Last Updated on STN: 20010322
Entered Medline: 20001121
AB Secretin is an endocrine hormone that stimulates the secretion of bicarbonate-rich pancreatic fluids. Recently, it has been discussed that secretin deficiency may be implicated in autistic syndrome, suggesting that the hormone could have a neuroendocrine function in addition to its role in digestion. In the present study, the human secretin gene (SCT) was isolated from a bacterial artificial chromosome genomic library. SCT contains four exons, with the protein coding regions spanning 713 bp of genomic DNA. Human SCT is similar structurally to the secretin genes of other species. Amino acid conservation, however, is most pronounced within the exon encoding the biologically active mature peptide. Northern blot analysis shows that human SCT transcripts are located in the spleen, intestinal tract, and brain. Radiation hybrid mapping places the SCT locus on chromosome 11p15.5.
Copyright 2000 S. Karger AG, Basel.

L7 ANSWER 4 OF 8 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
AN 1997:533418 BIOSIS
DN PREV199799832621
TI The effects of learning of phenytoin in adult rats exposed to the drug during development.
AU Besheer, J.; **Holloway, J. L.**; Banks, M. K.; Phipps, E.; Garraghty, P. E.
CS Dep. Psychol., Indiana Univ., Bloomington, IN 47405 USA
SO Society for Neuroscience Abstracts, (1997) Vol. 23, No. 1-2, pp. 2166.
Meeting Info.: 27th Annual Meeting of the Society for Neuroscience New Orleans, Louisiana, USA October 25-30, 1997
ISSN: 0190-5295.
DT Conference; Abstract; Conference
LA English

L7 ANSWER 5 OF 8 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
AN 1997:226709 BIOSIS
DN PREV199799518425
TI Sequence-based genetic markers for genes and gene families: Single-strand conformational polymorphisms for the fatty acid synthesis genes of Cuphea.
AU Slabaugh, M. B. (1); Huestis, G. M.; Leonard, J.; **Holloway, J. L.**; Rosato, C.; Hongtrakul, V.; Martini, M.; Toepfer, R.; Voetz, M.; Schell, J.; Knapp, S. J.
CS (1) Dep. Crop Soil Sci., Oregon State Univ., Corvallis, OR 97331 USA
SO Theoretical and Applied Genetics, (1997) Vol. 94, No. 3-4, pp. 400-408. ISSN: 0040-5752.
DT Article
LA English
AB Gene sequences are rapidly accumulating for many commercially and scientifically important plants. These resources create the basis for developing sequence-based markers for mapping and tracking known (candidate) genes, thereby increasing the utility of genetic maps. Members of most of the gene families underlying the synthesis of seed oil fatty acids have been cloned from the medium-chain oilseed Cuphea. Allele-specific-PCR (AS-PCR) and single-strand conformational polymorphism (SSCP) markers were developed for 22 fatty acid synthesis genes belonging to seven gene families of Cuphea using homologous and heterologous DNA sequences. Markers were developed for 4 fatty-acyl-acyl carrier protein thioesterase, 2 beta-ketoacyl-acyl carrier protein synthase I, 4 beta-keto-acylacyl carrier protein synthase II, 3 beta-ketoacyl-acyl carrier protein synthase III, 3 acyl carrier protein, 2 beta-ketoacyl-acyl carrier protein reductase, and 4 enoylacyl carrier protein reductase loci. Eighty-eight percent (14 of 16) of the SSCP loci were polymorphic, whereas only 9% (2 of 22) of the AS-PCR loci were polymorphic. These markers were mapped using a Cuphea viscosissima times C. lanceolata F-2 population and produced linkage groups of 10, 3, and 2 loci (3 loci segregated independently). The 10-locus linkage group had every gene but one necessary for the synthesis of 2- to 16-carbon fatty acids from acetyl-CoA and malonyl-ACP (the missing gene family was not mapped). SSCP analysis has broad utility for DNA fingerprinting and mapping genes and gene families.

L7 ANSWER 6 OF 8 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
AN 1995:392480 BIOSIS
DN PREV199598406780
TI Mapping dominant markers using F-2 matings.
AU Knapp, S. J. (1); **Holloway, J. L.**; Bridges, W. C.; Liu, B.-H.
CS (1) Dep. Crop Soil Sci., Oregon State Univ., Corvallis, OR 97331 USA
SO Theoretical and Applied Genetics, (1995) Vol. 91, No. 1, pp. 74-81. ISSN: 0040-5752.
DT Article
LA English
AB The development of efficient methods for amplifying random DNA sequences by the polymerase chain reaction has created the basis for mapping virtually unlimited numbers of mixed-phase dominant DNA markers in one population. Although dominant markers can be efficiently mapped using many different kinds of matings, recombination frequencies and locus orders are often mis-estimated from repulsion F-2 matings. The major problem with these matings, apart from excessive sampling errors of recombination frequency (theta) estimates, is the bias of the maximum-likelihood estimator (MLE) of theta (theta-ML). cxa theta-ML = 0 when the observed frequency of double-recessive phenotypes is 0 and the observed frequency of double-dominant phenotypes is less than 2/3 - the bias for those samples is -theta. We used simulation to estimate the mean bias of theta-ML. Mean bias is a function of n and theta and decreases as n increases. Valid maps of dominant markers can be built by using sub-sets of markers linked in coupling, thereby creating male and female coupling maps, as long as the maps are fairly dense (about 5 cM) - the sampling

errors of theta increase as theta increases for coupling linkages and are equal to those for backcross matings when theta = 0. The use of F-2 matings for mapping dominant markers is not necessarily proscribed because they yield twice as many useful markers as a backcross population, albeit in two maps, for the same number of DNA extractions and PCR assays; however, dominant markers can be more effeciently exploited by using doubled-haploid, recombinant-inbred, or other inbred populations.

L7 ANSWER 7 OF 8 MEDLINE
AN 96039020 MEDLINE
DN 96039020 PubMed ID: 7584391
TI Aligning genomes with inversions and swaps.
AU **Holloway J L**; Cull P
CS Oregon State University, USA.
SO ISMB, (1994) 2 195-202.
Journal code: 9509125.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199512
ED Entered STN: 19960124
Last Updated on STN: 19960124
Entered Medline: 19951214
AB The decision about what operators to allow and how to charge for these operations when aligning strings that arise in a biological context is the decision about what model of evolution to assume. Frequently the operators used to construct an alignment between biological sequences are limited to deletion, insertion, or replacement of a character or block of characters, but there is biological evidence for the evolutionary operations of exchanging the positions of two segments in a sequence and the replacement of a segment by its reversed complement. In this paper we describe a family of heuristics designed to compute alignments of biological sequences assuming a model of evolution with swaps and inversions. The heuristics will necessarily be approximate since the appropriate way to charge for the evolutionary events (delete, insert, substitute, swap, and invert) is not known. The paper concludes with a pairwise comparison of 20 Picornavirus genomes, and a detailed comparison of the hepatitis delta virus with the citrus exocortis viroid.

L7 ANSWER 8 OF 8 MEDLINE DUPLICATE 2
AN 91221915 MEDLINE
DN 91221915 PubMed ID: 2025147
TI Treatment for cigarette smoking in a Department of Veterans Affairs outpatient clinic.
AU Weissfeld J L; **Holloway J L**
CS Medical Service, Department of Veterans Affairs Medical Center, Ann Arbor, Michigan.
SO ARCHIVES OF INTERNAL MEDICINE, (1991 May) 151 (5) 973-7.
Journal code: 0372440. ISSN: 0003-9926.
CY United States
DT (CLINICAL TRIAL)
Journal; Article; (JOURNAL ARTICLE)
(RANDOMIZED CONTROLLED TRIAL)
LA English
FS Abridged Index Medicus Journals; Priority Journals
EM 199106
ED Entered STN: 19910623
Last Updated on STN: 19910623
Entered Medline: 19910606
AB A randomized clinical trial assessed the effectiveness of control, low-intensity, and high-intensity stop-smoking treatments in a Department of Veterans Affairs outpatient setting. The study actively recruited male cigarette smokers attending outpatient clinics at a university-affiliated

Veterans Affairs medical center. Subjects in the control group received an informational leaflet on smoking. Subjects in the low-intensity treatment group received a self-help booklet and a 20- to 30-minute session with a trained counselor. Subjects in the high-intensity group received the low-level treatments and individually tailored follow-up treatments provided in person, over the telephone, and through the mail. At least 6 months after randomization or last treatment, biochemically verified 1-month quit-smoking rates were 1.2% in 173 control subjects, 6.3% in 143 low-intensity treated subjects, and 6.0% in 150 high-intensity treated subjects. When rigorously defined, quit rates in each of the treated groups differed significantly from the control rate, but not from each other. The results demonstrated the effectiveness of moderately intensive stop-smoking treatments in a clinical setting of considerable interest, but not the incremental effectiveness of progressively more intensive treatments.

=>

---Logging off of STN---

=>
Executing the logoff script...

=> LOG Y

| COST IN U.S. DOLLARS | SINCE FILE ENTRY | TOTAL SESSION |
|---|---|---|
| FULL ESTIMATED COST | 22.33 | 22.54 |

STN INTERNATIONAL LOGOFF AT 15:36:00 ON 08 JUL 2003